

# Analysis of the Optimal Resource Allocation for a Tandem Queueing System

Liu Zaiming, Chen Gang, Wu Jinbiao\*

*School of Mathematics and Statistics, Central South University, Changsha 410083,  
Hunan, PR China*

---

## Abstract

In this paper, we study a controllable tandem queueing system consisting of two nodes and a controller, in which customers arrive according to a Poisson process and must receive service at both nodes before leaving the system. A decision maker dynamically allocates the number of service resource to each node facility according to the number of customers in each node. In the model, the objective is to minimize the long-run average costs. We cast these problems as Markov decision problems by dynamic programming approach and derive the monotonicity of the optimal allocation policy and the relationship between the two nodes' optimal policy. Furthermore, we get the conditions under which the optimal policy is unique and has the bang-bang control policy property.

*Keywords:*

Markov decision problem, Tandem system, Optimal policy, Dynamic programming, Average costs

---

## 1. Introduction

We consider a controllable tandem queueing system consisting of two nodes and a controller. A decision maker can assign a number of service resource to each node. The study of the controllable tandem queueing system is motivated by its wide applications in manufacturing, computer systems, voice and data communications, and vehicular traffic flow. The theory of

---

*Email addresses:* math\_lzm@csu.edu.cn (Liu Zaiming), chengmathcsu@163.com (Chen Gang), \* Corresponding author: wujinbiao@csu.edu.cn (Wu Jinbiao\*)

controllable queueing systems has often been studied for optimal control of admission, servicing, dynamic pricing, routing and scheduling of jobs in queues or networks of queues. These works are discussed in Stidham and Weber (1993), Yang et al. (2011) and Çil et al. (2011). The controllable queueing systems based on the theory of Markov, semi-Markov and regenerative decision processes can be found in Morozov and Steyaert (2013). Using the theory of the queueing system, we often cast the optimal problems as Markov decision problems (MDP). In order to get the properties of the optimal policy, the properties (such as the monotonicity, convexity property) of relative value function (when we consider the long-run average criteria) should be first considered. The key of the method is dynamic programming. For more details, we can see the paper written by Koole (1998) and Çil et al. (2009).

Based on the application background, the problems of the service resource control in different queueing systems have been investigated. Rykov and Efrosinin (2004) considered a multi-server controllable queueing system with heterogeneous servers, and several monotonicity properties of optimal policies are proved. Iravani et al. (2007) studied the optimal service scheduling in nonpreemptive finite-population queueing systems. The single-queue systems of the optimal resource allocation policy were considered by Yang et al. (2013). Efrosinin et al. (2014) analyzed a tandem queueing system of admission optimal policy.

Of particular relation to the present work are the works of Rosberg et al. (1982) and Ahn et al. (2002) where only the customer's holding cost was considered. Rosberg et al. (1982) considered the optimal control of service in tandem queues where the service rate in node 1 can be selected from a compact set and constant in node 2. Optimal control of a two-stage tandem queues system with flexible servers was discussed in Ahn et al. (2002) where only two flexible servers were considered under two different scenarios and they obtained the exhaustive optimal policy. Kaufman et al. (2005) considered the problem on the agile, temporary workforce into a tandem queueing system in which the relationship between the service rate and the number of the service resource is linear and the service resource costs in different nodes have the same cost function. However, different from the previous studies about resource allocation control problem, the two nodes in our model have the different holding cost rate and service resource cost function in the objective (long-run average cost). The main contribution of this paper is that we derive the monotonicity of the optimal allocation policy

and the relationship between the two nodes' optimal policy. Furthermore, we get the conditions under which the optimal policy is unique and the bang-bang control policy is established.

The rest of the paper is organized as follows. In Section 2, the model is formulated in detail based on the controllable Markov decision problem. The characteristics of the optimization problem and the optimality equation are derived in Section 3. In Section 4, structural properties of the optimal policy and main results of the paper are given. Finally, some further discussions and conclusions are given in Section 5.

## 2. Model Description

We consider a tandem queueing system with two nodes. Customers arrive at node 1 from outside the system according to a Poisson process with parameter  $\lambda$  and have exponentially distributed service requirement at each node. After receiving service at node 1, customers proceed immediately to node 2 and receive service before leaving the system. A decision maker can assign a number of service resource to each node. The service rate of a customer depends on the number of service resource assigned to the customer precisely. When a customer has been allocated  $a$  server resources, the service duration of that customer in node  $i$  is exponentially distributed with parameter  $\mu_i(a)$ ,  $i = 1, 2$ , which is strictly increasing in  $a$ . Without loss of generality, we assume that  $\mu_i(0) = 0$ ,  $i = 1, 2$ . At any decision epoch, the decision maker decides to choose the number of server resources to node 1 from a compact set  $A = [0, a_{max}]$ , and to node 2 from a compact set  $B = [0, b_{max}]$  at the same time. Each node has a single infinite-size FCFS queue. The interarrival and service times are assumed to be mutually independent. We assume that the stability condition  $\lambda < \mu_1(a_{max})$ ,  $\lambda < \mu_2(b_{max})$  holds. Figure 1 gives an illustration of the system.

We consider the following cost structure in the system. Our objective is to obtain dynamic management policy that minimizes the long-run average costs.

(1) resources cost: when the node  $i$  uses  $a$  resources, a cost of  $c_i(a)$ ,  $i = 1, 2$  is incurred by the system per unit time (here  $c_i(a)$  is a continuous function and strictly increasing in  $a$ . Without loss of generality, we assume that  $c_i(0) = 0$ ,  $i = 1, 2$ ).

(2) holding cost: holding costs are incurred at rate  $h_1$  and  $h_2$  per unit time for each customer in node 1 and 2, respectively.

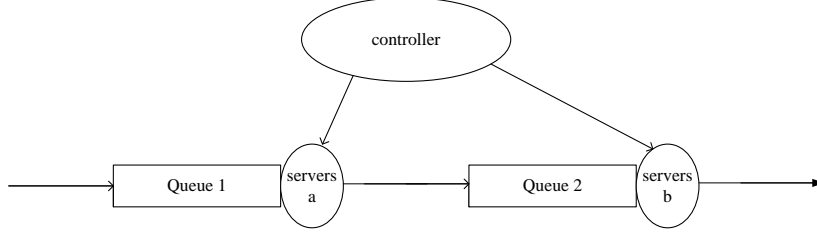


Fig. 1 The controllable tandem queueing systems

Let  $X_i(t)$  denote the number of customers at node  $i, i = 1, 2$ . The system evolves as a continuous-time Markov process

$$\{X(t), t \geq 0\} = \{(X_1(t), X_2(t)), t \geq 0\}.$$

The notations  $l_i(x), i = 1, 2$ , will be used to specify the certain components of the vector state  $x \in E$ .

The system state space is:  $E = x = (x_1, x_2) \in N^2$ , with  $N = 0, 1, 2, \dots$

It is assumed that the model is stable and conservative. The transition rate under a control action  $(a, b)$  is given by

$$Q_{xy}(a, b) = \begin{cases} \lambda & y = x + e_1; \\ \mu_1(a) & y = x - e_1 + e_2, l_1(x) > 0; \\ \mu_2(b) & y = x - e_2, l_2(x) > 0; \\ 0 & \text{else,} \end{cases}$$

where

$$Q_{xy}(a, b) \geq 0, y \neq x, Q_{xx}(a, b) = -Q_x(a, b) = -\sum_{y \neq x} Q_{xy}(a, b), Q_x(a, b) < \infty.$$

Here  $e_i$  is the 2-dimensional vector with 1 in the  $i$ th coordinate and 0 elsewhere,  $i = 1, 2$ .

The problem of the decision maker is to derive an optimal policy based on the number of customers in each node that minimizes the long-run average costs. We cast the customer resource management problem as a Markov decision problem. The set of decision epochs corresponds to the set of all arrivals, service completions, and dummy transitions due to uniformization. The controllable system associated with a Markov process is a five-tuple

$$\{E, D = (A, B), Q(f), c_i(a), h_i\}, i = 1, 2,$$

in which  $Q(f)$  is the transition matrix of the queueing system under the policy  $f$ .

We consider the stationary Markov policy  $f : E \rightarrow D$  with  $f = (f_1, f_2)$ . Due to the Markov property, it is clear that the optimal policy depends only on the current state regardless of  $t$ . More precisely, when the system state is  $x = (x_1, x_2)$ , the controller makes an action  $f_1(x_1) = a \in A, f_2(x_2) = b \in B$ . The action of the service resource to node  $i$  only depends on the current number of customers in node  $i$ .

### 3. Optimization problem and optimality equation

For every fixed stationary policy  $f$ , we assume that the process  $\{X(t), t \geq 0\}$  with state space  $E$  is an irreducible, positive recurrent Markov process. As it is known from Tijms (1994), for ergodic Markov process with the long-run average cost per unit of time for the policy  $f$  coincides with corresponding assemble average,

$$g(f) = \lim_{t \rightarrow \infty} u(x, t)^f / t = \sum_{i=1} \sum_{j=1} [c_1(f_1(i)) + c_2(f_2(j)) + h_1 i + h_2 j] \pi_{ij}(f), \quad (1)$$

in which  $u(x, t)^f$  denotes the total expected costs up to time  $t$  when the system starts in state  $x$  and  $\pi_{ij}(f)$  denotes a stationary probability of the process under policy  $f$ . The goal is to find a policy  $f^*$  that minimizes the long-term average costs:

$$g(f^*) = \min_f g(f). \quad (2)$$

In order to find the optimal policy  $f^*$  that minimizes the total average cost, we construct a discrete-time equivalent of the original system by using the standard tools of uniformization and normalization. Without loss of generality, we assume that  $\lambda + \mu_1(a_{max}) + \mu_2(b_{max}) = 1$ . Now we consider a real-valued function  $v(x)$  that plays the role of the relative value function, i.e., the asymptotic difference in total costs that results from starting the process in state  $x$  instead of some reference state. As it is well known, the optimal policy  $f$  and the optimal average cost  $g$  are the solutions of the optimality equation

$$Tv(x) = v(x) + g,$$

where  $T$  is the dynamic programming operator acting on  $v$ , defined as follows

$$Tv(x) = \lambda v(x + e_1) + \sum_{i=1,2} T_i v(x) + \sum_{i=1,2} h_i l_i(x), \quad (3)$$

here

$$T_1 v(x) = \min_{a \in A} \{ \mu_1(a) v(x - e_1 + e_2) + [\mu_1(a_{max}) - \mu_1(a)] v(x) + c_1(a) \}, \quad (4)$$

$$T_2 v(x) = \min_{b \in B} \{ \mu_2(b) v(x - e_2) + [\mu_2(b_{max}) - \mu_2(b)] v(x) + c_2(b) \}. \quad (5)$$

The first term in the expression  $Tv(x)$  models the arrivals of customers to node 1 from outside the system and the last one the customer holding cost. Similarly the first term in the expression  $T_1 v(x)$  corresponds to a customer who finished his service in node 1 and into node 2 and the second one the uniformization constant. The last one in  $T_1 v(x)$  is the resources cost in node 1. The first term in the expression  $T_2 v(x)$  corresponds to a customer who finished his service in node 2 and the second one the uniformization constant. The last one in  $T_2 v(x)$  is the resources cost in node 2.

According to (1), we can solve another optimization problem: if  $c_i \equiv 0, h_i = 1, i = 1, 2$ , then (2) is equivalent to minimization of the mean number of customers in the queueing system.

#### 4. Structural properties of the optimal policy

In this section, we focus on deriving the optimal policy. However, the optimal policy possesses structural properties that provide fundamental insight, and this also enables one to determine the optimal policy with less computational effort due to a reduction of the solution search space.

In order to study the structure, in principle, one needs to solve the optimal equation  $Tv(x) = v(x) + g$ . However it is hard to solve analytically in practice. It can be obtained by recursively defining  $v_{n+1} = Tv_n$  for arbitrary  $v_0$ . We know that the actions converge to the optimal policy as  $n \rightarrow \infty$ . For existence and convergence of the solutions and optimal policy we refer to Aviv and Federgruen (1999) and Sennott (2009). The backward recursion equation is given by

$$v_{n+1}(x) = \lambda v_n(x + e_1) + \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} h_i l_i(x).$$

For ease of notation, we define the set of the optimal policy in state  $x$  by:

$$f(x) = (f_1(x_1), f_2(x_2)) \quad f_1(x_1) = \arg T_1 v(x) \quad f_2(x_2) = \arg T_2 v(x).$$

By using the optimality equation, we can get the properties of relative value function as follows:

**Property 4.1 (non-decreasingness)**

- (i)  $v(x + e_i) \geq v(x)$ ,  $i = 1, 2$  for all  $x \in E$ ,
- (ii) if  $2h_2 \geq h_1$  then  $v(x - e_1 + e_2) \geq v(x - e_2)$  for all  $x = (x_1, x_2) \in E$  and  $x_1 \geq 1, x_2 \geq 1$ ,
- (iii) if  $h_1 \geq h_2$  then  $v(x) \geq v(x - e_1 + e_2)$  for all  $x = (x_1, x_2) \in E$  and  $x_1 \geq 1, x_2 \geq 1$ .

**Property 4.2 (quasi-convexity)**

- (i)  $v(x + e_2) - 2v(x) + v(x - e_2) \geq 0$ , for all  $x = (x_1, x_2) \in E$  and  $x_2 \geq 1$ ,
- (ii)  $v(x + e_1 - e_2) - 2v(x) + v(x - e_1 + e_2) \geq 0$ , for all  $x = (x_1, x_2) \in E$  and  $x_1 \geq 1, x_2 \geq 1$ .

Next we show some structure properties of the optimal policy, based on the structure properties of the relative value function above.

**Theorem 1.** *The optimal policy has the monotonicity property, i.e.,*

- (i) *if  $b_1 \in \arg T_2 v(x + e_2), b_2 \in \arg T_2 v(x)$ , then  $b_1 \geq b_2$  for all  $x = (x_1, x_2) \in E$ .*
- (ii) *if  $a_1 \in \arg T_1 v(x + e_1), a_2 \in \arg T_1 v(x)$ , then  $a_1 \geq a_2$  for all  $x = (x_1, x_2) \in E$ .*

The proof of Property 4.1 is given in Appendix A. The proof of Property 4.2 and Theorem 1 are given in Appendix B.

Based on Property 4.1, we give the relationship between the two nodes' optimal policy under some conditions.

**Theorem 2.** *Assume that  $c_1(a) - c_1(b) \geq c_2(a) - c_2(b)$  and  $\mu_2(a) - \mu_2(b) \geq \mu_1(a) - \mu_1(b)$  when  $a \geq b$ . Then if  $a \in \arg T_1 v(x), b \in \arg T_2 v(x)$ , we have  $b \geq a$  for all  $x = (x_1, x_2) \in E$  and  $x_1 \geq 1, x_2 \geq 1$ .*

PROOF. Let  $(a \in \arg T_1 v(x), b \in \arg T_2 v(x))$  be an arbitrary optimal policy for node 1 and 2 in state  $x$ , respectively. The proof is by contradiction. Suppose that  $b < a$ , then we compare the policy  $(a, b)$  with the policy  $(b, a)$ . We have:

$$\begin{aligned}
& T_{a,b} v_n(x) - T_{b,a} v_n(x) \\
&= [\mu_1(a)v(x - e_1 + e_2) + [\mu_1(a_{max}) - \mu_1(a)]v(x) + c_1(a)] \\
&\quad + [\mu_2(b)v(x - e_2) + [\mu_2(b_{max}) - \mu_2(b)]v(x) + c_2(b)] \\
&\quad - [\mu_1(b)v(x - e_1 + e_2) + [\mu_1(b_{max}) - \mu_1(b)]v(x) + c_1(b)] \\
&\quad - [\mu_2(a)v(x - e_2) + [\mu_2(a_{max}) - \mu_2(a)]v(x) + c_2(a)] \\
&= [\mu_1(a) - \mu_1(b)][v(x - e_1 + e_2) - v(x)] - [\mu_2(a) - \mu_2(b)][v(x - e_2) - v(x)] \\
&\quad + c_1(a) - c_1(b) - c_2(a) + c_2(b) \\
&\geq [\mu_1(a) - \mu_1(b)][v(x - e_1 + e_2) - v(x - e_2)] + c_1(a) - c_1(b) - c_2(a) + c_2(b) \\
&\geq 0.
\end{aligned}$$

The first equality is based on the definition of the operators  $T_1$  and  $T_2$ . The second equality follows by rearranging the terms. The first inequality follows the condition  $\mu_2(a) - \mu_2(b) \geq \mu_1(a) - \mu_1(b)$  when  $a \geq b$ . This implies that  $a$  and  $b$  is not an optimal policy for node 1 and 2 in state  $x$ , respectively. Hence,  $b \geq a$ .

From the above theorem we can conclude that under some conditions the optimal size of the service resources allocate to node 1 is less than that to node 2. We find that the optimal size of the resource allocate to each node depends on the resource cost variation  $c(a) - c(b)$  and the service rate variation  $\mu(a) - \mu(b)$  in each node.

We are now ready to give some conditions under which the optimal policy is unique and is a bang-bang control policy.

**Theorem 3.** *The following properties hold*

- (i) *if the functions  $m_1(a) = \frac{c'_1(a)}{\mu'_1(a)}$  and  $m_2(b) = \frac{c'_2(b)}{\mu'_2(b)}$  are monotonous on  $a \in A, b \in B$ , then the optimal policy is unique.*
- (ii)  *$\arg T_1 v(0) = \{0\}, \arg T_2 v(0) = \{0\}$ .*
- (iii) *if the functions  $\frac{c_1(a)}{\mu_1(a)}$  and  $\frac{c_2(b)}{\mu_2(b)}$  are non-increasing,  $\frac{c'_1(a)}{\mu'_1(a)} > \frac{c_1(a)}{\mu_1(a)}$  and  $\frac{c'_2(b)}{\mu'_2(b)} > \frac{c_2(b)}{\mu_2(b)}$  for all  $a \in (0, a_{max}), b \in (0, b_{max})$ , then the optimal policy is a bang-bang control policy. i.e.,  $\arg T_1 v(x) = \{0, a_{max}\}, \arg T_2 v(x) = \{0, b_{max}\}$  for all  $x \in E$ .*



PROOF. To prove part (i), we consider the optimal policy  $a$  in node 1 service resource allocation. In our event operator  $T_1$  for node 1 defined in equation (3), we have the following minimization problem:

$$T_1 v(x) = \min_{a \in A} \{ \mu_1(a) v(x - e_1 + e_2) + [\mu_1(a_{max}) - \mu_1(a)] v(x) + c_1(a) \}.$$

Rearranging the first-order optimality condition of the above problem, we have:

$$\frac{c'_1(a)}{\mu'_1(a)} = v(x) - v(x - e_1 + e_2).$$

Because the allocation resource action  $a \in A = [0, a_{max}]$ , the optimal policy  $a$  must be the solution of the above equation. Since the function  $m_1(a) = \frac{c'_1(a)}{\mu'_1(a)}$  is monotonous on  $a \in A$ , there is a unique  $a$  solving the above equation. Hence the optimal policy for node 1 is unique. The part (i) for node 2 can be proved in a similar manner.

To prove part (ii), we consider the optimal policy  $a$  in node 1 service resource allocation. As the problem is defined in equation (3), we have

$$T_1 v(0) = \min_{a \in A} \{ \mu_1(a) v(0) + [\mu_1(a_{max}) - \mu_1(a)] v(0) + c_1(a) \},$$

which immediately implies that  $\arg T_1 v(0) = \{0\}$ . The part (ii) for node 2 that  $\arg T_2 v(0) = \{0\}$  can be proved in a similar manner.

To prove part (iii), we consider the optimal policy  $a$  in node 1 service resource allocation. Since the service resources in node 1 is from the compact set  $[0, a_{max}]$ , the optimal policy  $a$  in node 1 can be 0, or  $a_{max}$ , or satisfies the following equation:

$$\frac{c'_1(a)}{\mu'_1(a)} = v(x) - v(x - e_1 + e_2).$$

We use the contradiction method. Assume that  $a \in \arg T_1 v(x)$  such that  $a \in (0, a_{max})$  for all  $x \in E$ . For any  $\varepsilon > 0$ , we have:

$$\begin{aligned} T_1^{a+\varepsilon} v(x) - T_1^a v(x) \\ = [\mu_1(a + \varepsilon) - \mu_1(a)] [v(x - e_1 + e_2) - v(x)] + c_1(a + \varepsilon) - c_1(a) \geq 0, \end{aligned}$$

which implies that

$$v(x) - v(x - e_1 + e_2) \leq \frac{c_1(a + \varepsilon) - c_1(a)}{\mu_1(a + \varepsilon) - \mu_1(a)}.$$

Since the function  $\frac{c_1(a)}{\mu_1(a)}$  is non-increasing, we get  $\frac{c_1(a+\varepsilon)-c_1(a)}{\mu_1(a+\varepsilon)-\mu_1(a)} \leq \frac{c_1(a)}{\mu_1(a)}$ ,  $v(x) - v(x - e_1 + e_2) \leq \frac{c_1(a)}{\mu_1(a)}$  which is a contradiction with the condition  $\frac{c'_1(a)}{\mu'_1(a)} > \frac{c_1(a)}{\mu_1(a)}$ . So there is no  $a$  satisfying the above equation. That is, the optimal policy in node 1 is  $\arg T_1 v(x) = \{0, a_{max}\}$ . Thus, the optimal policy is a bang-bang control policy. The part (iii) for node 2 can be proved in a similar manner.

## 5. Conclusion

In this paper we have analysed the optimal server resources control of a tandem queueing system with two nodes. The controller can make a dynamic decision to allocate the service resource to each node at any decision epoch. Applying the dynamic programming to the model, we not only give some traditional properties of the relative value function and optimal policy, but also derive the condition under which the optimal policy is unique and bang-bang control occurs. In particular, we have provided the relationship between the two nodes' optimal policy, which can give the controller more information to manage the system.

From the above results there arise some interesting extensions of the model which we may study in the near future.

(i) One possible change is to consider a model where each node's service resource decision is dependent on the number of the customers in two queues. When the system state is  $x = (x_1, x_2)$ , the controller makes an action  $f_1(x_1, x_2) = a \in A, f_2(x_1, x_2) = b \in B$ . Although the analysis is difficult, we may get some another properties of the queue optimal policy. In our model the two nodes have their action sets. We can also study the further model in which the two nodes share the common server resources.

(ii) Another way to generalize the model is to consider some strategies in our model, such as the retrial, feedback and priority customers. The model may become more complex. Some other methods should be considered. In our model the customers arrive at the system according to a Poisson process and the service time of a customer is exponentially distributed. We can apply the embedded Markov chain and semi-Markov decision processes to consider

the queueing system in which the service time of a customer is a general distribution.

(iii) In addition, the tandem queueing system with  $n$  nodes is also worthy thinking about. Based on our model, we can study the optimal policy relationship between the two nodes.

## Appendix A

Property 4.1 (non-decreasingness)

PROOF. To prove Property 4.1 (i), the proof is done by induction on  $n$  in  $v_n$ . Define  $v_0(x) = 0$  for all state  $x \in E$ . This function obviously satisfies (i). Now, we assume that (i) holds for the function  $v_n(x), x \in E$  and some  $n \in N$ . We should prove that  $v_{n+1}(x)$  satisfies the non-decreasing property as well. Then for  $i = 1$ , we can get

$$\begin{aligned} v_{n+1}(x + e_1) - v_{n+1}(x) \\ = \lambda[v_n(x + 2e_1) - v_n(x + e_1)] + h_1 + \sum_{i=1,2} T_i v_n(x + e_1) - \sum_{i=1,2} T_i v_n(x). \end{aligned}$$

The second term of the right-hand side is obviously positive.

Let  $(a \in \arg T_1 v(x), b \in \arg T_2 v_n(x))$  be an arbitrary optimal policy for node 1 and 2 in state  $x$ , respectively. Then

$$\begin{aligned} & \sum_{i=1,2} T_i v_n(x + e_1) - \sum_{i=1,2} T_i v_n(x) \\ & \geq \mu_1(a)[v_n(x + e_2) - v_n(x + e_2 - e_1)] \\ & \quad + \mu_2(b)[v_n(x - e_2 + e_1) - v_n(x - e_2)] \\ & \quad + [\mu_1(a_{max}) - \mu_1(a) + \mu_2(b_{max}) - \mu_2(b)][v_n(x + e_1) - v_n(x)] \\ & \geq 0, \end{aligned}$$

Therefore, Property 4.1 (i) holds by induction for any  $n$ ,  $v(x)$  is a nondecreasing function. Property 4.1 (i) for  $i = 2$  can be proved in a similar manner.

To prove Property 4.1 (ii), the proof is similar to the above one. Define  $v_0(x) = 0$  for all state  $x \in E$ . This function obviously satisfies the (ii). Now, we assume that (ii) holds for function  $v_n(x)$ ,  $x \in E$  and some  $n \in N$ . We should prove that  $v_{n+1}(x)$  satisfies Property 4.1 (ii) as well.

$$v_{n+1}(x - e_1 + e_2) - v_{n+1}(x - e_2)$$

$$\begin{aligned}
&= \lambda[v_n(x + e_2) - v_n(x + e_1 - e_2)] + 2h_2 - h_1 \\
&\quad + \sum_{i=1,2} T_i v_n(x - e_1 + e_2) - \sum_{i=1,2} T_i v_n(x - e_2).
\end{aligned}$$

Since the condition  $2h_2 \geq h_1$  holds, the second term of the right-hand side is obviously positive.

Let  $(a \in \arg T_1 v(x - e_2), b \in \arg T_2 v(x - e_2))$  be an arbitrary optimal policy for node 1 and 2 in state  $x - e_2$ , respectively. Then

$$\begin{aligned}
&\sum_{i=1,2} T_i v_n(x - e_1 + e_2) - \sum_{i=1,2} T_i v_n(x - e_2) \\
&\geq \mu_1(a)[v_n(x - 2e_1 + 2e_2) - v_n(x - e_2)] \\
&\quad + \mu_2(b)[v_n(x - e_1) - v_n(x - 2e_2)] \\
&\quad + [\mu_1(a_{\max}) - \mu_1(a)][v_n(x - e_1 + e_2) - v_n(x - e_2)] \\
&\quad + [\mu_2(b_{\max}) - \mu_2(b)][v_n(x - e_1 + e_2) - v_n(x - e_2)] \\
&\geq 0.
\end{aligned}$$

Therefore, Property 4.1 (ii) holds by induction for any  $n$ , we have  $v(x - e_1 + e_2) \geq v(x - e_2)$  for all  $x = (x_1, x_2) \in E$  and  $x_1 \geq 1, x_2 \geq 1$ . Property 4.1 (iii) can be proved in a similar manner.

## Appendix B

Property 4.2 (quasi-convexity) (i) and Theorem 1 (i)

PROOF. To prove Property 4.2 (i), we assume that Property 4.2 (i) for function  $v_n(x)$ ,  $x \in E$  and some  $n \in N$  holds. Then we need to prove that Property 4.2 (i) for  $n + 1$  also holds. When  $x = (x_1, x_2) \in E$  and  $x_2 \geq 1$ , we have

$$\begin{aligned}
&v_{n+1}(x + e_2) - 2v_{n+1}(x) + v_{n+1}(x - e_2) \\
&= \lambda[v_n(x + e_2 + e_1) - 2v_n(x + e_1) + v_n(x + e_1 - e_2)] \\
&\quad + \sum_{i=1,2} T_i v_n(x + e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_2) \\
&\geq \sum_{i=1,2} T_i v_n(x + e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_2).
\end{aligned}$$

The inequality holds by the induction hypothesis. The optimal policy of node 1 is only dependent on the number of customers in node 1 and the

state  $x + e_2$ ,  $x$ ,  $x - e_2$  have the same first entry  $x_1$ . Hence, they have the same optimal policy in node 1. We assume that  $a \in \arg T_1 v(x + e_2)$ ,  $b_1 \in \arg T_2 v(x + e_2)$ ,  $a \in \arg T_1 v(x - e_2)$ ,  $b_2 \in \arg T_2 v(x - e_2)$ . Therefore, we get

$$\begin{aligned}
& \sum_{i=1,2} T_i v_n(x + e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_2) \\
& \geq \mu_1(a)[v_n(x - e_1 + 2e_2) - 2v_n(x - e_1 + e_2) + v_n(x - e_1)] \\
& \quad + [\mu_1(a_{max}) - \mu_1(a)][v_n(x + e_2) - 2v_n(x) + v_n(x - e_2)] \\
& \quad + [\mu_2(b_1) - \mu_2(b_2)][v_n(x) - v_n(x - e_2)] \\
& \quad + \mu_2(b_2)[v_n(x) - 2v_n(x - e_2) + v_n(x - 2e_2)] \\
& \quad + [\mu_2(b_{max}) - \mu_2(b_1)][v_n(x + e_2) - v_n(x)] \\
& \quad + [\mu_2(b_{max}) - \mu_2(b_2)][v_n(x - e_2) - v_n(x)] \\
& = \mu_1(a)[v_n(x - e_1 + 2e_2) - 2v_n(x - e_1 + e_2) + v_n(x - e_1)] \\
& \quad + [\mu_1(a_{max}) - \mu_1(a)][v_n(x + e_2) - 2v_n(x) + v_n(x - e_2)] \\
& \quad + \mu_2(b_2)[v_n(x) - 2v_n(x - e_2) + v_n(x - 2e_2)] \\
& \quad + [\mu_2(b_{max}) - \mu_2(b_1)][v_n(x + e_2) - 2v_n(x) + v_n(x - e_2)] \\
& \geq 0.
\end{aligned}$$

The first inequality follows by taking a potentially suboptimal action in the second term of  $\sum_{i=1,2} T_i v_n(x + e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_2)$ . The equality follows by rearranging the terms. The last inequality follows by the induction hypothesis. Hence, we have  $v(x + e_2) - 2v(x) + v(x - e_2) \geq 0$ .

For Theorem 1 (i), let  $(b_1 \in \arg T_2 v(x + e_2), b_2 \in \arg T_2 v(x))$  be an optimal policy for node 2 in states  $x + e_2$ ,  $x$ , respectively. The proof is done by contradiction. Suppose that  $b_1 < b_2$ , then

$$\begin{aligned}
& T_2^{b_1} v(x) - T_2^{b_2} v(x) \\
& = [\mu_2(b_2) - \mu_2(b_1)][v(x) - v(x - e_2)] - [c_2(b_2) - c_2(b_1)] \geq 0.
\end{aligned}$$

Since Property 4.1 (i) above and  $\mu_2(b_2) - \mu_2(b_1) > 0$  holds, we have

$$\begin{aligned}
& T_2^{b_1} v(x + e_2) - T_2^{b_2} v(x + e_2) \\
& = [\mu_2(b_2) - \mu_2(b_1)][v(x + e_2) - v(x)] - [c_2(b_2) - c_2(b_1)] \\
& > [\mu_2(b_2) - \mu_2(b_1)][v(x) - v(x - e_2)] - [c_2(b_2) - c_2(b_1)] \\
& \geq 0.
\end{aligned}$$

However, this implies that  $b_1$  is not an optimal policy for node 2 in state  $x + e_2$ . Hence  $b_1 \geq b_2$ .

Property 4.2(quasi-convexity) (ii) and Theorem 1 (ii)

To prove Property 4.2 (ii), we assume that Property 4.2 (ii) holds for function  $v_n(x)$ ,  $x \in E$  and some  $n \in N$ . Then we need to prove that Property 4.2 (ii) for  $n + 1$  also holds. When  $x = (x_1, x_2) \in E$  and  $x_1 \geq 1, x_2 \geq 1$ , we have

$$\begin{aligned}
& v_{n+1}(x + e_1 - e_2) - 2v_{n+1}(x) + v_{n+1}(x - e_1 + e_2) \\
&= \lambda[v_n(x + 2e_1 - e_2) - 2v_n(x + e_1) + v_n(x + e_2)] \\
&\quad + \sum_{i=1,2} T_i v_n(x + e_1 - e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_1 + e_2) \\
&\geq \sum_{i=1,2} T_i v_n(x + e_1 - e_2) - 2 \sum_{i=1,2} T_i v_n(x) + \sum_{i=1,2} T_i v_n(x - e_1 + e_2) \\
&= T_1 v_n(x + e_1 - e_2) - 2T_1 v_n(x) + T_1 v_n(x - e_1 + e_2) \\
&\quad + T_2 v_n(x + e_1 - e_2) - 2T_2 v_n(x) + T_2 v_n(x - e_1 + e_2).
\end{aligned}$$

The inequality above holds by the induction hypothesis. Now, we assume that  $a_1 \in \arg T_1 v(x + e_1 - e_2)$ ,  $b_1 \in \arg T_2 v(x + e_1 - e_2)$ ,  $a_2 \in \arg T_1 v(x - e_1 + e_2)$ ,  $b_2 \in \arg T_2 v(x - e_1 + e_2)$ . Then, we get

$$\begin{aligned}
& T_1 v_n(x + e_1 - e_2) - 2T_1 v_n(x) + T_1 v_n(x - e_1 + e_2) \\
&\geq \mu_1(a_1)[v_n(x) - v_n(x - e_1 + e_2)] \\
&\quad + \mu_1(a_2)[v_n(x - 2e_1 + 2e_2) - v_n(x - e_1 + e_2)] \\
&\quad + [\mu_1(a_{max}) - \mu_1(a_1)][v_n(x + e_1 - e_2) - v_n(x)] \\
&\quad + [\mu_1(a_{max}) - \mu_1(a_2)][v_n(x - e_1 + e_2) - v_n(x)] \\
&= \mu_1(a_2)[v_n(x - 2e_1 + 2e_2) - 2v_n(x - e_1 + e_2) + v_n(x)] \\
&\quad + [\mu_1(a_{max}) - \mu_1(a_1)][v_n(x + e_1 - e_2) - 2v_n(x) + v_n(x - e_1 + e_2)] \\
&\geq 0.
\end{aligned}$$

The first inequality follows by taking a potentially suboptimal action in the second term of the operator  $T_1 v_n(x + e_1 - e_2) - 2T_1 v_n(x) + T_1 v_n(x - e_1 + e_2)$ . The equality follows by rearranging the terms. The last inequality follows by the induction hypothesis.

$$\begin{aligned}
& T_2 v_n(x + e_1 - e_2) - 2T_2 v_n(x) + T_2 v_n(x - e_1 + e_2) \\
&\geq \mu_2(b_1)[v_n(x + e_1 - 2e_2) - v_n(x - e_2)] \\
&\quad + \mu_2(b_2)[v_n(x - e_1) - v_n(x - e_2)]
\end{aligned}$$

$$\begin{aligned}
& +[\mu_2(b_{max}) - \mu_2(b_1)][v_n(x + e_1 + e_2) - v_n(x)] \\
& +[\mu_2(b_{max}) - \mu_2(b_2)][v_n(x - e_1 + e_2) - v_n(x)] \\
& = \mu_2(b_2)[v_n(x + e_1 - 2e_2) - 2v_n(x - e_2) + v_n(x - e_1)] \\
& +[\mu_2(b_{max}) - \mu_2(b_2)][v_n(x + e_1 - e_2) - 2v_n(x) + v_n(x - e_1 + e_2)] \\
& +[\mu_2(b_1) - \mu_2(b_2)][v_n(x + e_1 - 2e_2) - v_n(x + e_1 - e_2)] \\
& \geq 0.
\end{aligned}$$

The first inequality follows by taking a potentially suboptimal action in the second term of the operator above. The equality follows by rearranging the terms. The last one follows by the induction hypothesis and because of Theorem 1 (i), we know that  $b_1 \leq b_2$ . So that we have  $\mu_2(b_1) - \mu_2(b_2) \leq 0$ . From the Property 4.1, we know that  $v_n(x + e_1 - 2e_2) - v_n(x + e_1 - e_2) \leq 0$ . Thus, we derive that  $[\mu_2(b_1) - \mu_2(b_2)][v_n(x + e_1 - 2e_2) - v_n(x + e_1 - e_2)] \geq 0$ . Therefore, the last inequality is taken.

For Theorem 1 (ii), let  $(a_1 \in \arg T_1 v(x + e_1 - e_2), a_2 \in \arg T_1 v(x))$  be an optimal policy for node 2 in states  $x + e_1 - e_2, x$ , respectively. The proof is done by contradiction. Suppose that  $a_1 < a_2$ , then

$$\begin{aligned}
& T_1^{a_1} v(x) - T_1^{a_2} v(x) \\
& = [\mu_1(a_2) - \mu_1(a_1)][v(x - e_1 + e_2) - v(x)] - [c_1(a_2) - c_1(a_1)] \\
& \geq 0.
\end{aligned}$$

From Property 4.1 (ii) above and  $\mu_1(a_2) - \mu_1(a_1) > 0$ , we have

$$\begin{aligned}
& T_1^{a_1} v(x + e_1 - e_2) - T_1^{a_2} v(x + e_1 - e_2) \\
& = [\mu_1(a_2) - \mu_1(a_1)][v(x) - v(x + e_1 - e_2)] - [c_1(a_2) - c_1(a_1)] \\
& \geq [\mu_1(a_2) - \mu_1(a_1)][v(x - e_1 + e_2) - v(x)] - [c_1(a_2) - c_1(a_1)] \\
& \geq 0.
\end{aligned}$$

However, this implies that  $a_1$  is not an optimal policy for node 1 in state  $x + e_1 - e_2$ . Hence  $a_1 \geq a_2$ .

Since the optimal policy of node 1 is dependent only on the number of customers in node 1, and the states  $x + e_1, x + e_1 - e_2$  have the same first entry  $x_1 + 1$ . So they have the same optimal policy  $a_1$  in node 1, i.e.,  $a_1 \in \arg T_1 v(x + e_1)$ . Thus we get that if  $a_1 \in \arg T_1 v(x + e_1), a_2 \in \arg T_1 v(x)$  hold, then we have  $a_1 \geq a_2$  for all  $x = (x_1, x_2) \in E$ .

## References

- Ahn HS, Duenyas I, Lewis ME (2002) Optimal control of a two-stage tandem queueing system with flexible servers. *Probability in the Engineering and Informational Sciences* 16: 453–469.
- Aviv Y, Federgruen A (1999) The value iteration method for countable state markov decision processes. *Operations research letters* 24: 223–234.
- Çil EB, Karaesmen F, Örmeci EL (2011) Dynamic pricing and scheduling in a multi-class single-server queueing system. *Queueing Systems* 67: 305–331.
- Çil EB, Örmeci EL, Karaesmen F (2009) Effects of system parameters on the optimal policy structure in a class of queueing control problems. *Queueing Systems* 61: 273–304.
- Efrosinin D, Farhadov M, Kudubaeva S (2014) Performance analysis and monotone control of a tandem queueing system. In *Distributed Computer and Communication Networks*, 241–255. Springer.
- Iravani SM, Krishnamurthy V, Chao GH (2007) Optimal server scheduling in nonpreemptive finite-population queueing systems. *Queueing Systems* 55: 95–105.
- Kaufman DL, Ahn Hs, Lewis, ME (2005) On the introduction of an agile, temporary workforce into a tandem queueing system. *Queueing Systems* 51: 135–171.
- Koole G (1998) Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems* 30: 323–339.
- Morozov E, Steyaert B (2013) Stability analysis of a two-station cascade queueing network. *Annals of Operations Research* 202: 135–160.
- Rosberg Z, Varaiya PP, Walrand J (1982) Optimal control of service in tandem queues. *Automatic Control, IEEE Transactions on* 27: 600–610.
- Rykov V, Efrosinin D (2004) Optimal control of queueing systems with heterogeneous servers. *Queueing Systems* 46: 389–407.
- Sennott LI (2009) Stochastic dynamic programming and the control of queueing systems, vol. 504. John Wiley & Sons.
- Stidham Jr S, Weber R (1993) A survey of markov decision models for control of networks of queues. *Queueing systems* 13: 291–314.
- Tijms HC (1994) Stochastic models: an algorithmic approach, vol. 303. John Wiley & Sons Inc.
- Yang R, Bhulai S, Van der Mei R (2011) Optimal resource allocation for multiqueue systems with a shared server pool. *Queueing Systems* 68: 133–163.



Yang R, Bhulai S, van der Mei R (2013) Structural properties of the optimal resource allocation policy for single-queue systems. *Annals of Operations Research* 202: 211–233.